

A Real Time System for Hand Gesture Recognition

Prof Sonali N Jadhav¹
Professor, Dept. of Computer
Engineering
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail: khade.sonali@gmail.com

Mrunmayi Deshpande²
Student, Dept. of Computer Engineering
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail: mrun.desh@gmail.com

Atishay Jain³
Professor, Dept. of Computer
Engineering
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail: atishaycn@gmail.com

Shambhavi Gadhia⁴
Student, Dept. of Computer Engineering
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail: shambhavigadhia@gmail.com

Upendra Karki⁵
Student, Dept. of Computer Engineering
Rajiv Gandhi Institute of Technology
Mumbai, India
e-mail: upendrakarki28@gmail.com

Abstract—In this paper we explore the various aspects of hand gesture recognition in real time using neural networks. Hand gesture can be a vital way for the user to interact with any system. In this system we capture a hand gesture from the user and then perform the action related to it. This provides us with an alternative to mouse and keyboard to control a system. Hand gesture recognition can be helpful in various fields and areas where interacting with the system without touch is important. Hand gesture recognition is incorporated along with image processing and to add additional accuracy we are using neural network. This combination of image processing and neural network in real time forms a really powerful tool, forming the base of our project.

Keywords— *Hand gesture recognition, image processing, neural network, perceptron neural network*

I. INTRODUCTION

Communication in day to day life is vocal, but body language has its own significance. Hand gestures and facial expressions play an important role in conveying information. So gestures have many areas of application and can be used for various purposes.

Hand gesture recognition is a well-established idea with a great deal of potential, especially in fields where hands-free technology is in demand. Enabling touch free functionality to any existing system adds a completely new dimension to the existing application. Keeping this in mind we have discussed here gesture recognition with the help of computer science for accepting input, mathematical algorithms for matching and neural networks for training.

The process starts from accepting input from the user via a web camera. The input is stored as an image in the database. The image is then converted to a format that is machine readable, filtered and preprocessed. The feature extraction of this image is then done so as to compare the features of the input image with the images already available in the database. Once a match is found, the action coded corresponding to the matched image is performed. Here neural network has primary importance for training the system to produce accurate results. It is important that the system recognizes the input of the user and performs required task at highest possible level of accuracy. To achieve this, we train the system by giving it the same input a number of times. Once trained, the system has a low level of error and hence can be used in places where error margin needs to be very small

This report describes our project which is aimed at using neural networks particularly perceptron learning algorithm for making hand gesture recognition more accurate. Our aim here is to recognize the power of neural network along with image processing, making it a really strong technology. Our project is, in general, is extremely versatile as minor changes can be incorporated as per user requirement.

II. SYSTEM OVERVIEW

The hand gesture recognition system unfolds itself in two broad sub systems. The first being the calibration stage and the second, training. The first sub-system involves creating a detailed and valid database, accepting gestures from user and implementing the task as per commands. The second-sub system is completely dedicated to training, this is achieved the help of neural networks. We, in particular, are interested in the supervised neural network-perceptron neural network.

The working of the first sub system is extremely vast and thus, to make it simple let us divide it into four stages. The four stages are

1. Image acquisition
2. Pre-processing and gesture modelling
3. Feature extraction
4. Classification

2.1. Image Acquisition

Computer recognition of hand gestures may provide a more natural human computer interface, allowing people to point, or rotate by merely rotating their hands. We distinguish two

categories of gestures: static and dynamic. A static gesture is a particular hand configuration and pose, represented by a single image. A dynamic gesture is a moving gesture, represented by a sequence of images. We focus on the recognition of static gestures, although our method generalizes in a natural way to dynamic gestures. To understand a full message, it is necessary to interpret all the static and dynamic gestures over a period of time.

The first task of HGRS, the task deals with the capturing of still image with the help of the inbuilt java packages (JMF) and stores it in a predefined location in a prescribed format. Image capturing will be done continuously like video streaming. The model is matched to images of the hand by one or more cameras, and parameters corresponding to palm orientation and joint angles are estimated. These parameters are then used to perform gesture classification. To capture the image using a camera then extract some feature and those features are used as input in a classification algorithm for classification. If the hand signals fell in a predetermined set, and the camera views a close-up of the hand, we may use an example-based approach, combined with a simple method to analyze hand signals called orientation histograms. These example-based applications involve two phases; training and running. In the training phase, the user shows the system one or more examples of a specific hand shape. The computer forms and stores the corresponding orientation histograms. In the run phase, the computer compares the orientation histogram of the current image with each of the stored templates and selects the category of the closest match, or interpolates between templates, as appropriate. This method should be robust against small differences in the size of the hand but probably would be sensitive to changes in hand orientation.

2.2. Preprocessing & Gesture Modelling

The second module deals with processing the captured image. Image is being captured automatically once the system is booted as shown in fig 2. Once the image is captured then it is converted it into a gray scale image and then into a binary image so that gesture recognition task will be simplified. Filtering is being applied for removal of noise, 3x3 matrix filter is being designed for removal of noise from the binary image as shown in fig 3. Processing is a much required task in hand gesture recognition system. We have taken prima database which is standard database in gesture recognition. Images are processed before we can extract features from hand images. Preprocessing consist of two steps

2.2.1 Segmentation

Segmentation is done to convert gray scale image into binary image so that we can have only two object in image one is hand and other is background as shown in fig 1. Otsu algorithm is used for segmentation purpose and gray scale images are converted into binary image consisting hand or background as shown in fig 4. A very good segmentation is needed to select a adequate threshold of gray level for extract hand from background .i.e. there is no part of hand should have background and background also shouldn't have any part of hand. In general, the selection of an appropriate segmentation algorithm depends largely on the type of images and the application areas.

2.2.2 Morphological filtering

After converting gray scale image into binary image we have to make sure that there is no noise in image so we use morphological filter technique. Morphological techniques consist of four operations: dilation, erosion, opening and closing. Background may have some 1s which is known as background noise and hand gesture mat have some 0s that is known as gesture noise. These errors can lead to a problem in contour detection of hand gesture so we need to remove these errors. A morphological filtering approach has been applied using sequence of dilation and erosion to obtain a smooth, closed, and complete contour of a gesture.

2.3 Feature extraction

Once our input gesture as an image has been processed it is ready to be read and understood by the machine. The machine however, cannot read the gesture as it is. It is necessary that there is uniformity with respect to users. This is achieved with this module.

Specific features of the images are extracted and these are then used for comparison purposes by the computer. Since the image is already in binary form, there are a lot of methods to extract features. One can choose a method as per requirements. Some methods used the shape of the hand such as hand contour and silhouette while others utilized fingertips position, palm center, etc. created 13 parameters as a feature vector, the first parameters represents the ratio aspect of the bounding box of the hand and the rest 12 parameters are mean values of brightness pixels in the image. Self-Growing and Self-Organized Neural Gas (SGONG) neural algorithm is used to capture the shape of the hand, then three features are obtained; Palm region, Palm center, and Hand slope. This method is adopted by YCbCr color space extraction method. While another method, calculates the Center Of Gravity (COG) of the segmented hand and the distance from the COG to the farthest point in the fingers, and extracts one binary signal (1D) to estimate the number of fingers in the hand region. Various sign based methods use Gaussian pdf to extract geometric central moment as local and global features.



Fig 1. Background



Fig 2. Gesture



Fig 3. Subtraction



Fig 4. Grayscale

2.4 Classification

This is the last module and it deals with executing the particular associated operation with recognized hand gesture. If hand gesture recognized properly then predefined action will be invoked otherwise whole task will be repeated from the first module. Gesture classification method is used to recognize the gesture. Classification process is affected with the proper selection of feature parameters and classification algorithm. For example edge detection or contour operators cannot be used for gesture recognition since many hand postures are generated and this could produce misclassification. Euclidean distance metric used to classify the gestures for a real time system. Statistical tools used for gesture classification, HMM tool has shown its ability to recognize dynamic gestures such as drawing shapes in air. Besides, Finite State Machine (FSM), Learning Vector Quantization, and Principal Component Analysis (PCA). Neural network has been widely applied in the field of extracted the hand shape, and for hand gesture recognition. Other soft computing tools are effective in this field as well, such as Fuzzy Means clustering (FCM), and Genetic Algorithms (GA).

The second sub-system, training using neural network, is where accuracy of the proposed system comes into picture. The idea is that gestures of all types must be detected with minimal error. Neural networks can be categorized as: Supervised neural networks and unsupervised neural networks. In supervised, the output datasets are provided which are used to train the machine and get the desired outputs whereas in unsupervised no datasets are provided, instead the data is clustered into different classes. Since we have a database, we are interested in a supervised network. Thus, we have chosen perceptron.

When the system is being trained, the same input is given multiple times until it performs the right function for every possible variation in the gesture. Once the system generates right output all the time, the system is said to be trained and the training process is stopped. This process is called training

because we tell the system the desired output, give it the input and expect it to produce a output that is equivalent to desired output. If not, then the system is expected to learn from its mistake and keep processing the input till it produces a output equal to the desired output.

III. CONCLUSION

Hand gesture recognition provides us with a new way to interact with the system. This system can prove very useful in scenarios where use of various input devices is difficult. It can be helpful in medical fields to provide a sterile environment during surgeries, by allowing surgeons to operate different medical equipment without additional assistance. Hand gesture recognition can also assist the blind as they will be able to interact with system. With the fast computing power of today's technology it is possible to execute hand gesture actions in meaningful time. The process of capturing the image, processing it, recognizing the gesture and then performing the action associated with it can now be executed in a few seconds. If the response time of the performing gesture recognition in real time is not up to the mark then the entire idea of replacing the current system in various fields with more sophisticated hand gesture recognition will be in vain. Our proposed system aims at inculcating various technologies that would make the response time of such a system minimal. In addition to this, applying the concepts of neural networks makes the system cognitive. It ensures that the system gives the required output despite the agility of input. The system can be trained in the beginning but it also learns from every gesture successfully recognized. Thus, for a fast and accurate gesture recognition we have used neural networks.

ACKNOWLEDGMENT

We thank Professor Sonali N Jadhav, who has helped us in making the project. She has assisted us in every step. She has also shown a great deal of confidence in us while we were working on the project.

REFERENCES

- [1] V. S. Kulkarni, S.D.Lokhande, (2010) "Appearance Based Recognition of American Sign Language Using Gesture Segmentation", International Journal on Computer Science and Engineering (IJCSSE), Vol. 2(3), pp. 560-565.
- [2] Simej G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And NeuralInternational Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012
- [3] E. Stergiopoulou, N. Papamarkos. (2009). "Hand gesture recognition using a neural network shape fitting technique," Elsevier Engineering Applications of Artificial Intelligence, vol. 22(8), pp. 1141– 1158, doi: 10.1016/j.engappai.2009.03.008
- [4] Min B., Yoon, H., Soh, J., Yangc, Y., & Ejima, T. (1997). "Hand Gesture Recognition Using Hidden Markov Models". IEEE International Conference on computational cybernetics and simulation. Vol. 5, Doi: 10.1109/ICSMC.1997.637364
- [5] Malima, A., Özgür, E., Çetin, M. (2006). "A Fast Algorithm for Vision-Based Hand Gesture Recognition For Robot

-
- Control”, IEEE 14th conference on Signal Processing and Communications Applications, pp. 1-4. doi: 10.1109/SIU.2006.1659822
- [6] Simej G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). “A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural International Journal of Artificial Intelligence & Applications (IJAA), Vol.3, No.4, July 2012
- [7] M. M. Hasan, P. K. Mishra, (2011). “HSV Brightness Factor Matching for Gesture Recognition System”, International Journal of Image Processing (IJIP), Vol. 4(5).
- [8] Mokhar M. Hasan, Pramod K. Mishra, (2012) “Features Fitting using Multivariate Gaussian Distribution for Hand Gesture Recognition”, International Journal of Computer Science & Emerging Technologies IJCSET, Vol. 3(2).
- [9] Mokhar M. Hasan, Pramod K. Mishra, (2012). “Robust Gesture Recognition Using Gaussian Distribution for Features Fitting”, International Journal of Machine Learning and Computing, Vol. 2(3).